TITLE: **Reconciliation of Genetic and Genomic Approaches to Cotton Fiber Quality Improvement**

DISCIPLINE: **Genomics and Biotechnology**

AUTHORS:

J.-M. Lacape (Corresponding author)
CIRAD, UMR Développement et Amélioration des Plantes
TA A-96/03, Avenue Agropolis
34980 Montpellier Cedex 5, France
Phone: 33.(0)467614471
Fax: 33.(0)467615605
Email: marc.lacape@cirad.fr

M. Claverie
CIRAD, UMR Développement et Amélioration des Plantes
TA A-96/03, Avenue Agropolis
34980 Montpellier Cedex 5, France

J. Jacobs
Bayer BioScience N.V.
Technologiepark 38
B-9052 Gent
Belgium

D. Llewellyn
Y. Al-Ghazi
Elizabeth S. Dennis
CSIRO Plant Industry
P.O. Box 1600
Canberra ACT 2601
Australia

T. Arioli
Bayer BioScience N.V.
Technologiepark 38
B-9052 Gent
Belgium

R. Deryčker
Bayer BioScience N.V.
Technologiepark 38
B-9052 Gent
Belgium

N. Forestier-Chiron
CIRAD, UPR Systèmes Cotonniers
TA B-10/02, Avenue Agropolis
34980 Montpellier Cedex 5
France

M. Giband
CIRAD, UPR Systèmes Cotonniers
EMBRAPA-Algodão, Rua Oswaldo Cruz 1143
Centenário, 58.107-720
Campina Grande, PB, Brazil

J. Jean
CIRAD, UPR Systèmes Cotonniers
TA B-10/02, Avenue Agropolis
34980 Montpellier Cedex 5
France

C. Viot
CIRAD, UMR Développement et Amélioration des Plantes
TA A-96/03, Avenue Agropolis
34980 Montpellier Cedex 5
France

ACKNOWLEDGEMENT: This research is supported by the 2006 *Génoplante* programme of the French ANR (Agence Nationale de la Recherche).

ABBREVIATIONS: CIRAD, Centre de Coopération Internationale en Recherche Agronomique pour le Développement; CSIRO, Commonwealth Scientific and Industrial Research Organisation; AFLP, Amplified Fragment Length Polymorphism; RIL, Recombinant Inbred Line; PCR, Polymerase Chaine Reaction; Gb, Gigabases; cM, centimorgan

Reconciliation of Genetic and Genomic Approaches to Cotton Fiber Quality Improvement

ABSTRACT

The integration of genomics and plant breeding is driven by the increasing availability of sequence resources and by technological developments. The simultaneous measurement of the expression of thousands of genes is possible, and comparisons between contrasting genotypes and/or biological states, as well as within segregating populations has become feasible. In genetical genomics, the merger of genetics and genomics, gene expression profiles are quantitatively assessed within a segregating population, and expression quantitative trait loci (eQTL) can be mapped like classical QTLs. Methods and examples of applications related to genetical genomics will be reviewed, with emphasis on hybridisation-based (microarray) and PCR-based (cDNA-AFLP) techniques.

Despite the complexity of the molecular mechanisms underlying its development, the study of the cotton fiber has become a trait of primary interest. Several maps, including QTL maps, have been published, structural and metabolic genes related to fiber initiation or elongation have been identified, and several large EST projects have been developed. In this context, the applicability of a genetical genomics approach for the study of cotton fiber quality will be discussed.

A new cooperative project with CIRAD, Bayer CropScience and CSIRO, and supported by the French National Agency for Research, ANR, was initiated in 2007. The project aims at the genetic and genomic dissection of fiber quality using an interspecific *Gossypium hirsutum* X *G. barbadense* RIL population. Classical QTL mapping of fiber properties will be undertaken using data from different locations, and eQTLs will be detected using both microarray and cDNA-AFLP population-wide profiling.

KEY WORDS:

Cotton, eQTL, fiber, genetical genomics, mapping, QTL, recombinant inbred lines, RIL

Genomics breeding as an emerging discipline. Plant and animal breeders are facing the same major challenge, understanding the underlying causes of genetic variation of traits of interest. To this end, the methods and analytical tools relying on quantitative genetics used by the breeders have evolved over the last decade and the genomics revolution is now also having an impact on plant and animal improvement programmes (Varshney et al., 2005). In this context, the frontier between genetics and genomics weakens: as a discipline of biology, genetics is literally the study of heredity (the effect of genes on the phenotype), and genomics is the study of genes and genomes (their structure and function). By simplification one may consider that a "modern" molecular breeder, adopting a genomicist attitude must necessarily evolve from a more 'structuralist' point of view (constructing genetic maps, searching for quantitative trait loci and using marker-based selection) to a more 'functionalist' point of view (identifying key genes, their function and interrelations).

For the last 20 years (Paterson et al., 1988) so-called molecular breeding approaches, have essentially relied on genetic maps and on QTL detection. QTL mapping is used to identify chromosomal regions associated with a phenotypic trait of interest, i.e. displaying a correlation between the genotype (marker-derived) and the phenotype (trait value). Unfortunately, genome-wide scans with less than 200-300 individuals cannot resolve the location of a QTL more precisely than 10-30 cM (Schön et al., 2004). In the progression from QTL to gene, the more precise fine mapping of QTLs using larger recombinant populations, and subsequently the identification of candidate genes underlying those QTLs is conditioned by the availability of comprehensive sequence resources within the considered species or at least within sister species. Comparative genomics using sequence similarities with model organisms is offering a way to extend the gene repertoire of a given species. So far, the selection of quantitatively variable traits using classical phenotype-based QTL

strategies (marker-assisted selection) has met with limited success, and attempts to diagnose the true genetic bases of QTL are still scarce (Holland, 2007).

Large collections of expressed sequence tags (EST) are now available for most important plant species and the widespread use of high-throughput transcriptome sampling strategies (DNA chips), as well as the necessary bioinformatics developments, allow access to genome-wide gene expression profile data even for smaller crops of economic interest. Simultaneous measurement of the expression of thousands of genes in a trait-relevant tissue has been made possible. Typically one analyses patterns of expression that change over various treatments (such as mutant versus wild type, or induced versus uninduced) or over time.

The concept of Genetical-genomics. Jansen and Nap (2001) proposed the merger of genomics and genetics into the concept of genetical genomics (Figure): the expression levels of genes or cluster of genes are analysed within a segregating population. Basically, expression-QTLs (eQTL) controlling gene expression differences are mapped in a similar way as classical QTLs (Jansen and Nap, 2001; Schadt et al., 2003; de Koning and Haley, 2005). The approach provides a novel way of discovering, at a genetic level, regulators of gene expression acting either in *cis* or in *trans* relative to the target gene. Population- and genome-wide expression analysis also provides novel opportunities for correlating expression data to phenotypic/functional consequences. Cost-saving alternatives to large genome-wide population-wide analyses with minimal loss of informativeness have been proposed: analysing pooled samples of phenotypically extreme members of the population (Borevitz et al., 2003), or concentrating on genotypically-selected individuals (Xu et al., 2005).

The eQTL position may coincide with the gene itself displaying *cis* regulation (Borevitz and Chory, 2004; Kirst et al., 2005) or be different, thus revealing *trans*-acting factors controlling expression. A common feature of eQTL studies is the detection of

'hotspots' of *trans*-acting eQTL (Gibson and Weir, 2005; Keurenjes et al., 2007), interpreted as regions rich in regulatory genes that co-regulate many downstream targets.

Techniques to monitor gene expression rely either on hybridisation (microarrays) or on PCR (real-time PCR, cDNA-AFLP, Differential Display) or, more recently, on massive parallel sequencing (Solexa, Solid). Two main types of microarrays have been developed, namely cDNA-based and oligonucleotide-based arrays. Based on their technical specifications cDNA and oligonucleotide microarrays may differ in sensitivity and dynamic range for detecting variation in mRNA abundance as well as in power to discriminate between related target sequences. Unlike microarray technology, cDNA-AFLP is an open platform, not requiring prior sequence knowledge, and is also attractive for its low start-up cost (Vuylsteke et al., 2007). In quantitative cDNA-AFLP transcript profiling, AFLP gels are analysed to score band intensities. Differences in band intensity are expected to reflect fluctuations in transcript levels (Breyne et al., 2003). Variation in AFLP band intensity can be analysed for the study of regulatory responses through time and/or response to environment for any two contrasted genotypes (de Paepe et al., 2004) or apply to segregating populations (Vuylsteke et al., 2006).

Large-scale studies to map gene-expression differences (eQTLs) by individually profiling lines from a mapping population using either hybridisation-based or PCR-based detection methods, have been described in a relatively limited number of instances. Global eQTL mapping studies, using whole genome microarrays, have recently been published in yeast (Brem and Kruglyak 2005), using 112 segregants, and in *Arabidopsis* (West et al., 2007), using 211 RILs. Other successful applications at a more moderate scale of genetical genomics in plants include *Arabidopsis* (Keurenjes et al., 2007), maize (Schadt et al., 2003; Shi et al., 2007) and eucalyptus (Kirst et al., 2004).

Limitations and constraints of the different microarray platforms (Kuo et al., 2002, Mah et al., 2004) or in comparison with PCR-based techniques have been emphasized in several instances (Reijans et al., 2003; Tan et al., 2003). Questions to be addressed include: - optimising experimental designs, - accounting for technical and biological variation (intra- and inter-sample), - cross-hybridization between related genes and its interpretation, etc. The principal limitation for the microarray technique is that only a fraction of genes, those for which DNA sequence are available, can be investigated. Recently, the modelling strategies for expression data have been discussed by Pérez-Enciso et al. (2007): gene expressions being interconnected, an eQTL model for a given gene should consider all other genes (variables) as regressors (in a step-wise regression strategy). As a result, the significance of eQTL hotspots reported in some publications needs to be revisited.

In brief, a combination of different profiling methods is likely to be most informative (Jansen and Nap, 2001).

Cotton fiber quality as a target trait. As compared to other important crops, cotton faces some specific challenges due to its polyploid genome structure (the two major cultivated species are allotetraploid), the large genome size (2.3 Gb, 26 chromosomes and 5700 cM), the low level of polymorphism within the cultivated species and the complexity of fiber quality as a primary trait of interest. However, recent years have seen significant efforts aimed at a better understanding of the cotton genome in its structural organization, as well as of the cotton fiber transcriptome in its functional aspects.

The densest genetic maps of cotton are derived from interspecific crosses between *Gossypium hirsutum* and *G. barbadense* (Rong et al., 2004; Nguyen et al., 2004; Frelichowski et al., 2006; Han et al., 2006; Yu et al., 2007) as these two species display a workable level of molecular polymorphism and show interesting phenotypic variation. Some of these mapping populations have also served for QTL analyses, with emphasis on various

fiber quality traits (Lacape et al., 2005; Chee et al., 2005a; Chee et al., 2005b; Draye et al., 2005; Frelichowski et al., 2006; He et al., 2007). Validation of 'conserved' fiber quality QTLs across populations has not been conclusive, due to the fact that the majority of these QTL studies were either derived from small and mortal [F_2 or backcross (BCs)] populations, or were not replicated.

As compared to F_2 or BCs, homozygous F_2-immortalized recombinant inbred lines, RILs, constitute the preferred material for QTL mapping in many crops. RILs have not been widely utilized in cotton, mainly due to long development timelines and difficulties producing sufficient seed. Intraspecific RIL populations have been used to screen for markers linked to nematode resistance (Wang and Roberts, 2006) and to fiber quality (Wang et al., 2006; Shen et al., 2007). An interspecific RIL population derived from a cross between the cytogenetic standards of *G. hirsutum* and *G. barbadense* (TM1 and 3-79, respectively) has been used for genetic map construction (Park et al., 2005) and QTL analyses (Frelichowski et al., 2006).

A cotton fiber is a long single elongated cell of the ovule epidermis, averaging 25 to 29 mm in the most cultivated species (*G. hirsutum*) and reaching 50 mm in some *G. barbadense* accessions. The development of the cotton fiber is typically divided into four distinct, but overlapping stages: initiation when the number of fibers is determined, elongation when length is determined, secondary wall synthesis when strength is determined and maturation when fiber structure is determined. The cotton fiber transcriptome has attracted a lot of attention in recent years (see for example a review by Wilkins and Arpat (2005) and T. Wilkins at this conference). Many fiber-specific genes involved in fiber cell initiation, fiber elongation or cell wall biogenesis have been identified from the comparisons of normal (wild-type) versus fiber mutants of *G. hirsutum* species. Only few reports have investigated the mechanisms and genes underlying the important developmental differences

between *G. hirsutum* and *G. barbadense* (Ruan et al., 2004; Wu et al., 2005; Zhang et al., 2005).

An increasing number of sequence resources (BACs and ESTs) in *Gossypium* have been used to define markers providing landmarks, mostly as SSRs, on the genetic maps (http://www.mainlab.clemson.edu/cmd/) and several groups have designed fiber cDNA microarrays for functional studies (Zhang et al., 2005; Wu et al., 2006). From a total of over 30 EST libraries (>200 000 ESTs), Udall et al. (2006) identified a collection of ~33,000 exemplar sequences, and recent efforts from the cotton community lead to the public release of 2 cotton microrarrays (essentially from fiber ESTs), including a 24 K GeneChip® Cotton Genome Array from Affymetrix (http://www.affymetrix.com/products/arrays/specific/cotton.affx); and a 23 K oligonucleotide (60-70mer) microarray by Udall et al. (2007). Also important are the recent efforts from an international panel of scientists promoting the sequencing of the cotton genome(s) (see white paper "Cotton Genome Sequencing" available from http://icgi.tamu.edu/ and AH Paterson at this conference).

Despite this wealth of efforts to decipher the molecular determinants for the quality of the cotton fiber, this scientific community remains rather compartmentalized between functional and structural genomics research specialists. Though expression-QTL (eQTL) mapping is rapidly gaining recognition as a valuable approach for closing the gap between (structural) genetics and (functional) genomics in a range of organisms, it has never been applied to cotton.

Our project on the genetics and genomics of cotton fiber quality. The French National Agency for Research, ANR, has become the major channel for supporting both public and public-private research in France. A three year (2007-2009) project named Cotton_RILs has recently been granted 700 000 € by the ANR through its plant genomics specialized program, Génoplante. The project brings together three important participants in

the field of cotton molecular breeding and fiber genomics, CIRAD, Bayer CropScience and CSIRO. This project proposes to integrate genetic and genomic approaches for the characterization of cotton fiber quality using a new interspecific cotton RIL population created by CIRAD (Montpellier, France). The RIL population distinguishes itself by the excellent agronomic and fiber properties of the *G. hirsutum* and *G. barbadense* parents, and by the availability of a detailed collection of reference data (a high-density map, extensive QTL data) from genetically related backcross populations.

The project has 4 major components: (1) genetic map construction using AFLP and SSR markers (a saturated map of over 1000 loci is expected), (2) extensive evaluation of fiber properties in replicated trials on four continents (meta QTL-analysis), (3) genome-wide and population-wide expression analysis (eQTL mapping) for critical phase(s) of fiber development using two complementary profiling technologies, microarray and cDNA-AFLP, and (4) fine mapping of a limited set of fiber QTL and eQTL loci (candidate gene identification).

To our knowledge, the activities proposed in this project will be the first attempt to utilize an integrated approach of genetics and genomics (genetical genomics) for the characterisation of a cotton population. The compilation through meta-analysis of fiber QTL data from this study with data from the literature (the majority of the markers used are cross referenced in other populations), and the integration of QTL data with expression data (eQTL) are expected to help to identify chromosomal regions important for fiber quality as well as important candidate genes influencing fiber quality, and ultimately facilitate the breeding of superior genotypes.

CONCLUSION

The genomics, -post-genomics, revolution will be at the core of plant breeding only when it can elucidate the relationship between variation in phenotypic traits and the variation in gene sequences and expression (Morgante and Salamini, 2003). To this end, there is a need for integration of disciplines such as structural genomics, transcriptomics, proteomics with plant breeding and plant physiology implying an increasing need for a bioinformatics integration of datasets (Varshney et al., 2005).

Our study of an interspecific cotton RIL population, extensively studied over different sites both for their phenotype and the expression level of an important number of fiber transcripts, is expected to validate the concept of genetical genomics. This will constitute the first attempt to bridge structural and functional approaches to provide a better understanding of cotton fiber quality determination.

REFERENCES

Borevitz, J.O., and J. Chory. 2004. Genomics tools for QTL analysis and gene discovery. Curr. Opin. Plant Biol. 7: 132-136.

Borevitz, J.O., D. Liang, D. Plouffe, H. Chang, T. Zhu, D. Weigel, C.C. Berry, E. Winzeler, and J. Chory. 2003. Large-scale identification of single-feature polymorphisms in complex genomes. Genome Res. 13: 513-523.

Brem, R.B., and L. Kruglyak. 2005. The landscape of genetic complexity across 5700 gene expression traits in yeast. Proc. Natl. Acad. Sci. USA 102(5): 1572-1577.

Breyne, P., R. Dreesen, B. Cannoot, D. Rombaut, K. Vandepoele, S. Rombauts, R. Vanderhaeghen, Z. Inzé, and M. Zabeau. 2003. Quantitative cDNA-AFLP analysis for genome-wide expression studies. Mol. Genet. Genomics 269: 173-179.

Chee, P., X. Draye, C. X. Jiang, L. Decanini, T.A. Delmonte, R. Bredhauer, C.W. Smith, and A.H. Paterson. 2005a. Molecular dissection of phenotypic variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: III. Fiber length. Theor. Appl. Genet. 111: 772-781.

Chee, P., X. Draye, C.X. Jiang, L. Decanini, T.A. Delmonte, R. Bredhauer, C.W. Smith, and A.H. Paterson. 2005b. Molecular dissection of interspecific variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: I. Fiber elongation. Theor. Appl. Genet. 111: 757-763.

de Koning, D.J., and C.S. Haley. 2005. Genetical genomics in humans and model organisms. Trends Genet. 21: 377-381.

de Paepe, A., M. Vuylsteke, P. Van Hummelen, M. Zabeau, and D. Van der Straeten. 2004. Transcriptional profiling by cDNA-AFLP and microarray analysis reveals novel insights into the early response to ethylene in *Arabidopsis*. The Plant J. 39: 537-559.

Draye, X., P. Chee, C.X. Jiang, L. Decanini, T.A. Delmonte, R. Bredhauer, C.W. Smith, and A.H. Paterson. 2005. Molecular dissection of interspecific variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: II. Fiber fineness. Theor. Appl. Genet. 111: 764-771.

Frelichowski, J.E., M.B. Palmer, D. Main, J.P. Tomkins, R.G. Cantrell, D.M. Stelly, J. Yu, R.J. Kohel, and M. Ulloa. 2006. Cotton genome mapping with new microsatellites from Acala 'Maxxa' BAC-ends. Mol. Genet. Genomics 275: 479-491.

Gibson, G., and B. Weir. 2005. The quantitative genetics of transcription. Trends Genet. 21: 616-623.

Han, Z.G., C. Wang, X.L. Song, W. Guo, J.Y. Gou, C. Li, X. Chen, and T. Zhang. 2006. Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSRs in allotetraploid cotton. Theor. Appl. Genet. 112: 430-439.

He, D.H., Z. Lin, X. Zhang, Y.C. Nie, X.P. Guo, Y. Zhang, and W. Li. 2007. QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of *Gossypium hirsutum* x *Gossypium barbadense*. Euphytica 153: 181-197.

Holland, J.B. 2007. Genetic architecture of complex traits in plants. Curr. Opin. Plant Biol. 2: 156-161.

Jansen, R.C., and J.P. Nap. 2001. Genetical genomics: the added value from segregation. Trends Genet. 7: 388-391.

Keurenjes, J.J.B., J. Fu, I.R. Terpstra, J.M. Garcia, G. van den Ackerveken, L.B. Snoek, A.J.M. Peeters, D. Vreugdenhill, M. Koorneef, and R.C. Jansen. 2007. Regulatory network construction in *Arabidopsis* by using genome-wide gene expression quantitative trait loci. Proc. Natl. Acad. Sci. USA 104: 1708-1713.

Kirst, M., C.J. Basten, A.A. Myburg, Z.B. Zeng, and R.R. Sederoff. 2005. Genetic architecture of transcript-level variation in differentiating xylem of a *Eucalyptus* hybrid. Genetics 169: 2295-2303.

Kirst, M., A.A. Myburg, J.P.G. De Leon, M.E. Kirst, J. Scott, and R.R. Sederoff. 2004. Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of *Eucalyptus*. Plant Physiol. 135: 2368-2378.

Kuo, W.P., T.K. Jenssen, A.J. Butte, L. Ohno-Machado, and I.S. Kohane. 2002. Analysis of matched mRNA measurements from two different microarray technologies. Bioinformatics 18: 405-412.

Lacape, J.-M., T.B. Nguyen, B. Courtois, J.-L. Belot, M. Giband, J.-P. Gourlot, G. Gawryziak, S. Roques, and B. Hau. 2005. QTL analysis of cotton fiber quality using multiple *G. hirsutum* x *G. barbadense* backcross generations. Crop Sci. 45: 123-140.

Mah, N., A. Thelin, T. Lu, S. Nikolaus, T. Kühbacher, Y. Gurbuz, H. Eickhoff, G. Klöppel, H. Lehrach, B. Mellgard, C.M. Costelio, and S. Schreiber. 2004. A comparison of oligonucleotide and cDNA-based microarray systems. Physiol. Genomics 16: 361-370.

Morgante, M., and F. Salamini. 2003. From plant genomics to breeding practices. Curr. Opin. Biotech. 14: 214-219.

Nguyen, T.B., M. Giband, P. Brottier, A.-M. Risterucci, and J.-M. Lacape. 2004. Wide coverage of tetraploid cotton genome using newly developed microsatellite markers. Theor. Appl. Genet. 109: 167-175.

Park, Y.-H., M.S. Alabady, B. Sickler, T.A. Wilkins, J. Yu, D.M. Stelly, R.J. Kohel, O.M. El-Shihy, R.G. Cantrell, and M. Ulloa. 2005. Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line (RIL) cotton population. Mol. Genet. Genomics 274: 428-441.

Paterson, A.H., E. Lander, J.D. Hewitt, S. Peterson, S. Lincoln, and S. D. Tanksley. 1988. Resolution of quantitative traits into Mendelian factors using a complete linkage map of restriction fragment length polymorphism. Nature 335(20): 721-726.

Pérez-Enciso, M., J.R. Quevedo, and A. Bahamonde. 2007. Genetical genomics: use all data. BMC Genomics 8.

Reijans, M., R. Lascaris, A.O. Groeneger, A. Wittenberg, E. Wesselink, J. van Oeveren, E. de Wit, A. Boorsma, B. Voetdijk, H. van der Spek, L.A. Grivell, and G. Simons. 2003. Quantitative comparison of cDNA-AFLP, microarrays, and GeneChip expression data in *Saccharomyces cerevisiae*. Genomics 6: 606-618.

Rong, J., C. Abbey, J.E. Bowers, C.L. Brubaker, C. Chang, P. Chee, T.A. Delmonte, X. Ding, J.J. Garza, B.S. Marler, C.-h. Park, G.J. Pierce, K.M. Rainey, V.K. Rastogi, S.R. Schulze, N. Trolinder, J.F. Wendel, T.A. Wilkins, T.D. Williams-Coplin, R.A. Wing, R.J. Wright, X. Zhao, L. Zhu, and A.H. Paterson. 2004. A 3347 locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). Genetics 166: 389-417.

Ruan, Y.L., S.M. Xu, R. White, and R.T. Furbank. 2004. Genotypic and developmental evidence for the role of plasmodesmatal regulation in cotton fiber elongation mediated by callose turnover. Plant Physiol. 136: 4104-4113.

Schadt, E.E., S.A. Monks, D.A. Drake, A.L. Lusis, N. Che, V. Colinayo, T.G. Ruff, S.B. Milligan, J.R. Lamb, G. Cavet, P.S. Linsley, M. Mao, R.B. Stoughton, and S.H. Friend. 2003. Genetics of gene expression surveyed in maize, mouse and human. Nature 422: 297-302.

Schön, C., H.F. Utz, S. Groh, B. Truberg, S. Openshaw, and A.E. Melchinger. 2004. Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. Genetics 167: 485-498.

Shen, X., W. Guo, Q. Lu, X. Zhu, Y. Yuan, and T. Zhang. 2007. Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. Euphytica adv.online.

Shi, C., A. Uzarowska, M. Ouzunova, M. Landbeck, G. Wenzel, and T. Luebberstedt. 2007. Identification of candidate genes associated with cell wall digestibility and eQTL (expression quantitative trait loci) analysis in a Flint x Flint maize recombinant inbred line population. BMC Genomics 1:22.

Tan, P.K., T.J. Downey, E.L. Sptiznagel, P. Xu, D. Fu, D.S. Dimitrov, R.A. Lempicki, B.M. Raaka, and M.C. Cam. 2003. Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Res. 31: 5676-5684.

Udall, J.A., L.E. Flagel, F. Cheung, A.W. Woodward, R. Hovav, R.A. Rapp, J. Swanson, J.J. Lee, A.R. Gingle, D. Nettleton, C.D. Town, Z.J. Chen, and J.F. Wendel. 2007. Spotted cotton oligonucleotide microarrays for gene expression analysis. BMC Genomics 8: adv online.

Udall, J.A., J. Swanson, K. Haller, R.A. Rapp, M.E. Sparks, J. Hatfield, Y. Yu, Y. Wu, A.A.B, B.A. Sickler, T.A. Wilkins, J.-Y. Guo, J.-Y. Chen, J. Scheffler, E. Taliercio, R. Turley, H. Mc Fadden, P. Payton, R. Allen, D. Zhang, C. Haigler, C.G. Wilkerson, J. Suo, S.R. Schulze, M. Pierce, M. Essenberg, H. Kim, D. Llewellyn, E.S. Dennis, D. Kudma, R.A. Wing, A.H. Paterson, C.A. Soderlund, and J.F. Wendel. 2006. A global assembly of cotton ESTs. Genome Res. 16: 441-450.

Varshney, R.K., A. Graner, and M.E. Sorrells. 2005. Genomics-assisted breeding for crop improvement. Trends Plant Sc. 10: 621-630.

Vuylsteke, M., H. Daele, A. Vercauteren, M. Zabeau, and M. Kuiper. 2006. Genetic dissection of transcriptional regulation by cDNA-AFLP. Plant J. 3: 439-446.

Vuylsteke, M., J. Peleman, and M.J.T. van Eijk. 2007. AFLP-based transcript profiling (cDNA-AFLP) for genome-wide expression analysis. Nature Protocols 2 No. 6: 1399-1413.

Wang, B.-H., W. Guo, X. Zhu, Y. Wu, N. Huang, and T. Zhang. 2006. QTL mapping of fiber quality in an elite hybrid derived-RIL population of upland cotton. Euphytica 152: 367-378.

Wang, C., and P.A. Roberts. 2006. Development of AFLP and derived CAPS markers for root-knot nematode resistance in cotton. Euphytica 152: 185-196.

West, M.A., K. Kim, D.J. Kliebenstein, H. van Leeuwen, R.W. Michelmore, R.W. Doerge, and D.A. St. Clair. 2007. Global eQTL mapping reveals the complex genetic architecture of transcript level variation in *Arabidopsis*. Genetics 175 (3): 1441-1450.

Wilkins, T.A., and A. Bulak Arpat. 2005. The cotton fiber transcriptome. Physiol. Plantarum 124: 295-300.

Wu, Y., A.C. Machado, R.G. White, D. Llewellyn, and E.S. Dennis. 2006. Expression profiling identifies genes expressed early during lint fibre initiation in cotton. Plant Cell Physiol. 47(1): 107-127.

Wu, Z., K.M. Soliman, A. Zipf, S. Saha, G.C. Sharma, and J.N. Jenkins. 2005. Isolation and characterization of genes differentially expressed in fiber of *Gossypium barbadense* L. J. Cotton Sci. 9: 166-174.

Xu, Z., F. Zou, and T. Vision. 2005. Improving quantitative trait loci mapping resolution in experimental crosses by the use of genotypically selected samples. Genetics 170: 401-408.

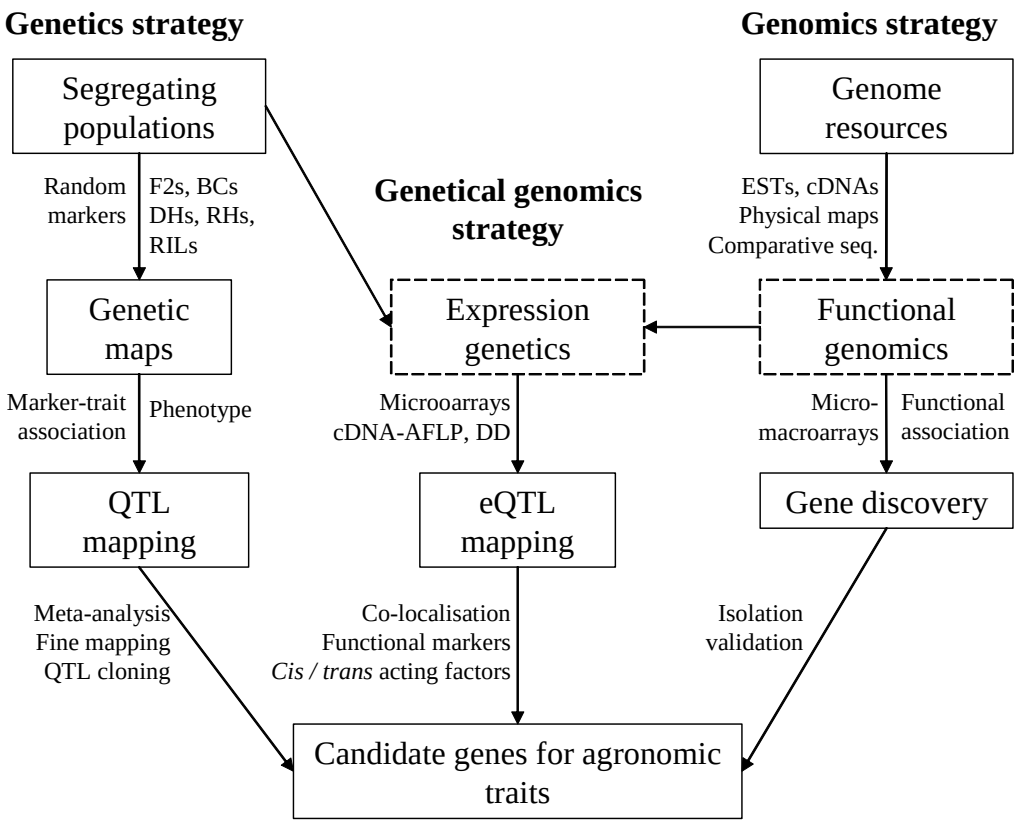
Yu, J., S. Yu, C. Lu, W. Wang, S. Fan, M. Song, Z. Lin, X. Zhang, and J. Zhang. 2007. High-density linkage map of cultivated allotetraploid cotton based on SSR, TRAP, SRAP and AFLP markers. J. Integr. Plant Biol. 49(5): 716-724.

Zhang, J., D.J. Hinchliffe, T.A. Wilkins, and R.G. Cantrell. 2005. Comparative microarray analysis of genes differentially expressed during fiber development of Upland and Pima cotton. Abstract. p. 883. *In* Proc. Beltwide Cotton Conf., New Orleans, LA. Natl. Cotton Counc. Amer., Memphis, TN.

FIGURE CAPTIONS

Figure 1. Genetical genomics strategy, combining genetics with gene expression analysis.

Fig. 1.


Genetics strategy
Genomics strategy
Segregating populations
Genome resources
Random markers
F2s, BCs DHs, RHs, RILs
Genetical genomics strategy
ESTs, cDNAs Physical maps Comparative seq.
Genetic maps
Expression genetics
Functional genomics
Marker-trait association
Phenotype
Microoarrays cDNA-AFLP, DD
Micro-macroarrays
Functional association
QTL mapping
eQTL mapping
Gene discovery
Meta-analysis Fine mapping QTL cloning
Co-localisation Functional markers *Cis / trans* acting factors
Isolation validation
Candidate genes for agronomic traits